**Table 7** Summary of Early⋆, Sub-flow†-based and Encrypted‡ flow feature-based traffic classification

| Ref. | ML Technique | Dataset | Features | Classes | Evaluation | |
|---|---|---|---|---|---|---|
| | | | | | Settings | Results |
| Bernaille et al. [55]* | Unsupervised *k*-Means | Proprietary: univ. network | Packet size and direction of first *P* packets in a flow | eDonkey, FTP, HTTP, Kazaa, NNTP, POP3, SMTP, SSH, HTTPS, POP3S | $P = 5, k = 50$ | Accuracy > 80% |
| TIE [108, 121]* | Supervised J48 DT, *k*-NN, Random Tree, RIPPER, MLP, NB | Proprietary: Univ. Napoli campus network | Payload size stats and inter-packet time stats of first *N* packets, bidirectional flow duration and size, transport protocol | BitTorrent, SMTP, Skype2Skype, POP, HTTP, SOULSEEK, NBNS, QQ, DNS,SSL RTP, EDONKEY | $N = 1...10$ | Overall accuracy = 98.4% with BKS (J48, Random Tree, RIPPER, PL) combiner, $N = 10$ |
| Nguyen et al. [337]† | Supervised NB, C4.5 DT | Proprietary: home network, univ. network, game server | Inter-packet arrival time statistics, inter-packet length variation statistics, IP packet length statistics of *N* consecutive packets | Enemy Territory (online game), VoIP, Other | $N = 25$ | C4.5 DT: Enemy Territory - recall* = 99.3%, prec.* = 97%; VoIP - recall* = 95.7%, precision* = 99.2%<br>NB: Enemy Territory - recall* = 98.9%, prec.* = 87%, VoIP - recall* = 99.6%, precision* = 95.4%<br>* median |
| Erman et al. [137]⋆ | Semi-supervised *k*-Means | Proprietary: Univ. Calgary | Number of packets, average packet size, total bytes, total header bytes, total payload bytes (caller to callee and *vice versa*) | P2P, HTTP, CHAT, EMAIL, FTP, STREAMING, OTHER | $k = 400$, 13 layers, packet milestones (number of packets) in layers are separated exponentially (8, 16, 32, …) | Flow accuracy > 94%, byte accuracy 70-90% |
| Li et al. [270]⋆ | Supervised C4.5 DT, C4.5 DT with AdaBoost, NBKE | Proprietary | A subset of 12 from 248 features [321] of first *N* packets | WEB, MAIL, BULK, Attack, P2P, DB, Service, Interactive | $N = 5$ | C4.5 DT: Accuracy >99%; Attack is an exception with moderate-high recall |
| Jin et al.[222]⋆ | Supervised AdaBoost | Proprietary: ISP network, labeled as in [176] | Lowsrcport, highsrcport, duration, mean packet size, mean packet rate, toscount, tcpflags, dstinnet, lowdstport, highdstport, packet, byte, tos, numtosbytes, srcinnet | Business, chat, DNS, FileSharing, FTP, Games, Mail, Multimedia, NetNews, SecurityThreat, VoIP, Web | Number of binary classifiers (*k*): TCP = 12, UDP = 8 | Error rate:<br>TCP = 3%, UDP = 0.4% |
| Bonfiglio et al. [69]‡ | Supervised NB, Pearson's $\chi^2$ test | Proprietary: univ. network, ISP network | Message size, average inter-packet gap | Skype | NB decision threshold $B_{min} = -5$, $\chi^2(Thr) = 150$ | NB∧$\chi^2$:<br>UDP – E2E - FP = 0.01%, FN = 29.98%<br>E2O - FP = 0.0%, FN = 9.82% (univ. dataset);<br>E2E - FP = 0.01%, FN = 24.62%<br>E2O - FP = 0.11%, FN = 2.40% (ISP dataset)<br>TCP – negligible FP |
| Alshammari et al. [17]‡ | Supervised AdaBoost, SVM, NB, RIPPER, C4.5 DT | AMP [457], MAWI [474], DARPA99 [278], proprietary from Univ. Dalhousie | Packet size, packet inter-arrival time, number of packets, number of bytes, flow duration, protocol (forward and backward direction) | SSH, Skype | N/A | C4.5 DT:<br>SSH – DR = 95.9%, FPR = 2.8% (Dalhousie), DR = 97.2%, FPR = 0.8% (AMP), DR = 82.9%, FPR = 0.5% (MAWI)<br>Skype – DR = 98.4%, FPR = 7.8% (Dalhousie) |

**Table 7** Summary of Early⋆, Sub-flow†-based and Encrypted‡ flow feature-based traffic classification *(Continued)*

| Ref. | ML Technique | Dataset | Features | Classes | Evaluation | |
|---|---|---|---|---|---|---|
| | | | | | Settings | Results |
| Shbair et al. [409]‡ | Supervised C4.5 DT, RF | Synthetic trace | Statistical features from encrypted payload and [253] (client to server and *vice versa*) | Service Provider (number of services): Uni-lorraine.fr (15), Google.com (29), akamihd.net (6), Googlevideo.com (1), Twitter.com (3), Youtube.com (1), Facebook.com (4), Yahoo.com (19), Cloudfront.com (1) | N/A | RF (service provider): precision = 92.6%, recall = 92.8%, F-measure = 92.6% RF (service): accuracy in 95-100% for majority of service providers > 100 connections per HTTPS service |

N/A: Not available

offsets in binary and textual protocols, such as DNS and HTTP, respectively. Though, the CSG resulted in a higher misclassification error, approximately 7%, it performed best for SSH encrypted traffic. However, it is important to realize that encryption often introduces randomness in the payload. Hence, techniques such as in Ma et al. [286] that search for keywords at fixed offsets will suffer in performance.

Techniques that rely on capturing the beginning of flows [176, 286] are infeasible for links with high data rates where sampling is often employed. Finamore et al. [146] overcome this limitation by extracting signatures

**Table 8** Summary of NFV⋆ and SDN†-based traffic classification

| Ref. | ML Technique | Dataset | Features | Classes | Evaluation | |
|---|---|---|---|---|---|---|
| | | | | | Settings | Results |
| He et al. [182]⋆ | Supervised *k*-NN, Linear-SVM, Radial-SVM, DT, RF, Extended Tree, AdaBoost, Gradient-AdaBoost, NB, MLP | KDD [42] | Protocol, network service, source bytes, destination bytes, login status, error rate, connection counts, connection percentages (different services among the same host, different hosts among the same service) | Attack types from [450] | Dynamic selection of classifier and features to collect | Accuracy = 95.6% |
| Amaral et al. [19]† | Supervised RF, SGBoost, XGBoost | Proprietary: enterprise network | Packet size (1 to *N* packets), packet timestamp (1 to *N* packets), inter-arrival time (*N* packets), source/destination MAC, source/destination IP, source/destination port, flow duration, packet count byte count | BitTorrent, Dropbox, Facebook, Web Browsing (HTTP), LinkedIn, Skype, Vimeo, YouTube | $N = 5$ | RF: Accuracy 73.6-96.0% SGBoost: Accuracy 71.2-93.6% XGBoost: Accuracy 73.6-95.2% |
| Wang et al. [462]† | Semi-supervised Laplacian-SVM | Proprietary: univ. network | Entropy of packet length, average packet length (source to destination and *vice versa*), source port, destination port, packets to respond from source to destination, minimum length of packets from destination to source, packet inactivity degree from source to destination, median of packet length from source to destination for the first *N* packets | Voice/video conference, streaming, bulk data transfer, interactive | $N = 20$, Laplacian-SVM parameters $\lambda = 0.00001 - 0.0001, \sigma = 0.21 - 0.23$ | Accuracy > 90% |

N/A: Not available